

<https://eventos.utfpr.edu.br//sei/sei2019>

## Avaliação da certificação CoreTrustSeal para o Repositório Brasileiro Livre para Dados Abertos do Solo

### Evaluation of the CoreTrustSeal certification for the Free Brazilian Repository for Open Data of Soil

#### RESUMO

**Matheus Ferreira Ramos**  
[matheusramos@alunos.utfpr.edu.br](mailto:matheusramos@alunos.utfpr.edu.br)  
Universidade Tecnológica Federal do Paraná, Santa Helena, Paraná, Brasil

**Alessandro Samuel-Rosa**  
[alessandrorosa@utfpr.edu.br](mailto:alessandrorosa@utfpr.edu.br)  
Universidade Tecnológica Federal do Paraná, Santa Helena, Paraná, Brasil

O volume de informação digital tem crescido a cada ano e essas informações precisam de meios práticos e seguros para serem armazenadas, utilizadas e reutilizadas. E assim, foram criados os repositórios digitais. Contudo, um problema que ocorre é que parte dos repositórios ainda não segue padrões mais amplamente aceitos para armazenamento e acesso aos dados. Isso dificulta o acesso e acarreta o potencial esquecimento dos repositórios e dos dados neles depositados. Pensando nisso, foram criadas certificações para repositórios de dados como o CoreTrustSeal. Essas certificações avaliam, entre outros aspectos, o grau de padronização, a confiança dos dados armazenados no repositório e a sua longevidade. Este trabalho tem como objetivo avaliar se o Repositório Brasileiro Livre para Dados Abertos do Solo (febr) atende aos requisitos necessários para obter a certificação CoreTrustSeal. Inicialmente, fez-se um levantamento dos requisitos para certificação CoreTrustSeal. Em seguida, analisou-se se o febr se enquadra em cada requisito analisado, atribuindo-lhe níveis de aderência. Os resultados mostraram que o febr necessita de ajustes para se enquadrar no padrão internacional CoreTrustSeal. Em especial, sugere-se a adoção de um esquema de metadados com padrões aceitos internacionalmente, trazendo assim mais segurança e confiabilidade para o repositório.

**PALAVRAS-CHAVE:** Repositório digital. Dados da pesquisa. Pedometria.

**Recebido:** 19 ago. 2019.

**Aprovado:** 01 out. 2019.

**Direito autorial:** Este trabalho está licenciado sob os termos da Licença Creative Commons-Atribuição 4.0 Internacional.



#### ABSTRACT

The volume of digital information has been growing every year, and this digital information needs practical and secure means to be stored, used and reused. And so the digital repositories were created. However, a problem that occurs is that some of the repositories still do not follow more widely accepted standards for data storage and access. This makes it difficult to access data and leads to the potential forgetting of repositories and the data stored therein. With that in mind, certifications have been created for data repositories like CoreTrustSeal. These certifications assess, among other things, the degree of standardization, the reliability of the data stored in the repository and its longevity. This work aims to evaluate if the Free Brazilian Repository for Open Soil Data (febr) meets the necessary requirements to obtain the CoreTrustSeal certification. First, CoreTrustSeal certification requirements were surveyed. Then, it was analyzed if the febr fits in each analyzed requirement, giving it adherence levels. The results showed that febr needs adjustments to fit the CoreTrustSeal international standard. In particular, it is suggested to adopt a metadata scheme with internationally accepted standards, thus bringing more security and reliability to the repository.

**KEYWORDS:** Digital repository. Research data. Pedometrics.

**2** FOME ZERO E AGRICULTURA SUSTENTÁVEL



**13** AÇÃO CONTRA A MUDANÇA GLOBAL DO CLIMA



**15** VIDA TERRESTRE



## INTRODUÇÃO

O volume de informação digital tem crescido a cada ano e essas informações digitais precisam de meios práticos e seguros para serem armazenadas, utilizadas e reutilizadas. Um bom sistema para fazer esse trabalho são os repositórios digitais, coleções que capturam e preservam a produção intelectual de uma comunidade (MASSON, 2008). Atualmente, existem muitos repositórios de dados armazenando todos os tipos de dados, como os institucionais, multidisciplinares, orientados a projetos de pesquisa e disciplinares (SAYÃO; SALES, 2016). Os repositórios institucionais de dados de pesquisas são caracterizados por possuírem vínculo acadêmico, os quais armazenam informações provenientes de pesquisas e trabalhos realizados na universidade ou instituto. Já os repositórios multidisciplinares são aqueles que podem manter dados de diversas áreas de conhecimento. Repositórios institucionais normalmente são multidisciplinares pois, em sua grande maioria, tratam de assuntos diferenciados, como por exemplo o Repositório Institucional da Universidade Tecnológica Federal do Paraná (RIUT). Existem também repositórios de dados da pesquisa, que são orientados ao armazenamento de dados produzidos em um projeto de pesquisa específico ou vários projetos desenvolvidos para resolver um problema específico. Por fim, os repositórios disciplinares, que são focados em dados específicos de uma área do conhecimento bem delimitada.

Um problema que ocorre é que parte dos repositórios de dados ainda não segue padrões amplamente aceitos para armazenamento e disponibilidade aos dados. Isso dificulta o acesso aos mesmos pela comunidade externa e demais usuários potenciais, e acarreta no potencial esquecimento do repositório e dos dados nele depositados. Pensando nisso, foram criadas certificações para estes a fim de padronizar, garantir que todos os dados armazenados nele sejam de confiança e que haja também longevidade do mesmo. Um dos sistemas de certificação, criado no fim de 2017 pela World Data System (WDS) of the International Science Council (ISC) em parceria com Data Seal of Approval (DSA), é o CoreTrustSeal (CTS). Essa certificação é baseada no modelo de auditoria DSA-WDS Core Trustworthy Data Repositories Requirements que foi baseado no modelo ISO-16363, o CTS é uma certificação de nível básico e tem 16 requisitos.

O Repositório Brasileiro Livre para Dados Abertos do Solo, febr, é o maior repositório de dados da pesquisa no Brasil. Desenvolvido como projeto de extensão tecnológica da Universidade Tecnológica Federal do Paraná, o febr tem como propósito servir como plataforma central para a compilação e armazenamento de dados do solo. No febr, os dados ficam disponíveis para qualquer usuário interessado e podem ser utilizados para qualquer finalidade. Para aumentar ainda mais a sua credibilidade e confiança junto à comunidade científica nacional e internacional, um dos próximos passos do projeto é conseguir a certificação CoreTrustSeal. O objetivo deste estudo é avaliar se o febr atende aos requisitos para obtenção da certificação CoreTrustSeal e propor medidas corretivas e melhorias em caso contrário.

## MATERIAL E MÉTODOS

Iniciou-se pelo levantamento e estudo detalhado de cada um dos 16 requisitos do CoreTrustSeal, conforme descrito na Tabela 1.

Tabela 1. Requisitos para certificação de repositórios de dados pelo CoreTrustSeal.

Item	Descrição
1	Tem uma missão explícita para fornecer acesso e preservar dados em seu domínio?
2	Mantém todas as licenças aplicáveis cobrindo o acesso a dados, o uso e monitora a conformidade?
3	Tem um plano de continuidade para garantir o acesso contínuo e a preservação de suas propriedades?
4	Garante que os dados sejam criados, curados, acessados e usados em conformidade com as normas disciplinares e éticas?
5	Tem financiamento adequado e um número suficiente de funcionários qualificados para realizar a missão com eficiência?
6	Adota mecanismo(s) para garantir orientação e feedback de especialistas em andamento ?
7	Garante a integridade e autenticidade dos dados?
8	Aceita dados e metadados com base em critérios definidos para garantir relevância e capacidade de compreensão para usuários de dados?
9	Aplica processos e procedimentos documentados no gerenciamento do armazenamento de arquivamento dos dados?
10	Assume a responsabilidade pela preservação de longo prazo e gerência essa função de forma planejada e documentada?
11	Tem experiência para abordar dados técnicos e qualidade de metadados e assegura que informações estejam disponíveis para os usuários finais?
12	O arquivamento ocorre de acordo com fluxos de trabalho definidos, desde a ingestão até a disseminação?
13	Permite que os usuários descubram os dados e consultem-nos de maneira persistente por meio de citação apropriada?
14	Permite a reutilização dos dados ao longo do tempo, garantindo que metadados estejam disponíveis para suportar seu entendimento e uso?
15	Funciona em sistemas operacionais bem suportados e usa tecnologias de hardware e software apropriadas aos serviços que fornece?
16	A infra-estrutura técnica fornece proteção das instalações e seus dados, produtos, serviços e usuários?

Fonte: DSA-WDS Core Trustworthy Data Repositories Requirements (2016).

Depois do estudo detalhado dos 16 requisitos do CoreTrustSeal, analisou-se o nível de aderência do febr a cada dos requisitos aplicáveis. Os níveis de

aderência utilizados foram os seguintes: atende, atende parcialmente, e não atende. Para o febr receber a certificação CoreTrustSeal, é preciso que o mesmo atenda de maneira completa a todos os requisitos aplicáveis. Assim, depois da avaliação, para cada requisito cuja aderência foi parcial ou nula, foram apontadas medidas corretivas e de melhoria do repositório.

### RESULTADOS E DISCUSSÃO

Dos 16 requisitos aplicáveis avaliados, o febr atende completamente a apenas três e de maneira parcial a outros sete. Cinco requisitos não são atendidos. A Tabela 2 apresenta o nível de aderência do febr aos diferentes requisitos e uma descrição da situação atual.

Tabela 2. Aderência do febr aos requisitos para certificação CoreTrustSeal.

Item	Situação	Avaliação
1	Não atende	O febr não tem em seu portal da web uma missão clara e explícita para fornecer e preservar os dados.
2	Atende parcialmente	A licença de uso dos dados está disponível em todos os conjunto de dados. Porém, não há um documento tratando sobre o descumprimento das condições de uso.
3	Não atende	Não se encontra preparativos para garantir o acesso aos dados em caso de mudança de circunstâncias.
4	Não aplicável	O febr não lida com dados sensíveis à divulgação.
5	Atende parcialmente	O febr tem financiamento para exercer suas funções, porém este não é garantido, pois depende de programas governamentais. Seus contribuidores detêm vasta experiência, além de estar buscando obter mais conhecimento em novas áreas.
6	Atende	Em seu portal pode-se encontrar um redirecionamento para um fórum onde é destinado a sugestões, críticas e suporte técnico.
7	Atende	Os dados são conferidos com os originais e o acesso de edição e alteração são restritos.
8	Atende	Para este requisito pode-se notar que o febr apresenta com antecedência seus padrões para que o depositante consiga entender como os dados devem ser depositados e para que o febr garanta a qualidade dos dados.
9	Não atende	Não foi constatado a existência de backups por parte do repositório, logo o mesmo não atendeu ao requisito.
10	Não atende	Não são encontrados documentos demonstrando que assume a responsabilidade sobre a preservação dos dados.

Fonte: Autoria própria (2019).

Tabela 2. Aderência do febr aos requisitos para certificação CoreTrustSeal. (Continuação)

Item	Situação	Avaliação
11	Atende parcialmente	A equipe é qualificada e consegue abordar termos técnicos, porém ainda não há metadados.
12	Atende parcialmente	Pode-se encontrar em seu web-site a documentação detalhada para o processo de ingestão, porém não consta nenhum documento para a disseminação.
13	Atende parcialmente	O febr apresenta algumas ferramentas de pesquisa em seu portal, porém ainda não contém os metadados nos padrões internacionais.
14	Não atende	Faltam metadados com padrões internacionais.
15	Atende parcialmente	O febr utiliza tecnologia de terceiros instalada em servidores na nuvem, mas possui pouco controle sobre essa tecnologia e os servidores.
16	Atende parcialmente	Como há dependência de tecnologia de terceiros, o febr oferece apenas proteção limitada.

Dentre os requisitos não atendidos, têm destaque aqueles relacionados à infraestrutura tecnológica, como o uso de metadados com padrões internacionais, a existência de backups, e um plano de contingência. Dos requisitos atendidos apenas parcialmente, a maioria refere-se à documentação e aos recursos humanos e financeiros disponíveis. Note-se que a resolução das questões relacionadas à infraestrutura tecnológica passam pela maior disponibilidade de recursos humanos e financeiros. Esses resultados mostram que o febr não está pronto para obter a certificação emitida pelo CoreTrustSeal.

O uso de repositórios de dados da pesquisa como o febr pode trazer consigo inúmeras vantagens para a ciência criando meios de compartilhamento e acesso a informações (COSTA; LEITE, 2010). Resultando em benefícios também para as instituições que os adotam, pois melhoram a sua visibilidade, aumentando, assim, sua comunicação com a comunidade interna e externa. Por fim, as vantagens chegam até os pesquisadores, que obtêm mais prestígio e maior visibilidade de suas produções. A lista de benefícios é ampliada contando também com preservação digital, segurança, curadoria e reuso dos dados, serviços inovadores e indicadores de qualidade e produtividade da instituição (SAYÃO; SALES, 2016). Para que todos esses potenciais do febr sejam efetivados, é preciso realizar correções e melhorias. A mais fácil de implementar e que se recomenda que seja abordada em primeiro lugar é a melhoria da documentação. Em segundo lugar, aconselha-se que haja maior investimento na melhoria da infraestrutura tecnológica. Como isso demandará maior volume de recursos financeiros e pessoal capacitado, cada vez mais escassos, recomenda-se que sejam estudadas maneiras alternativas de realizar tais correções e melhorias.

Uma delas é a maior atuação de alunos de cursos de graduação da área da ciência da computação.

### CONCLUSÃO

Este trabalho mostrou que, para obter a certificação CoreTrustSeal, ainda são necessários ajustes e melhorias no febr. As principais razões são a insuficiente documentação e disponibilidade de recursos financeiros e humanos, e a limitada infraestrutura tecnológica disponível.

A recomendação é que as correções e melhorias iniciem pelas deficiências de mais fácil resolução. Para aquelas mais complexas, que exigem maior disponibilidade de recursos, tanto humanos como financeiros, sugere-se que soluções alternativas mais baratas sejam buscadas.

### REFERÊNCIAS

- CORETRUSTSEAL. **Core Trustworthy Data Repositories Requirements**. v01.00. 2016. Disponível em:  
<[https://www.coretrustseal.org/wp-content/uploads/2017/01/Core\\_Trustworthy\\_Data\\_Repositories\\_Requirements\\_01\\_00.pdf](https://www.coretrustseal.org/wp-content/uploads/2017/01/Core_Trustworthy_Data_Repositories_Requirements_01_00.pdf)> acesso em: 17 de agosto 2019
- COSTA, Sely Maria de Souza; LEITE, Fernando César Lima. **Insumos conceituais e práticos para iniciativas de repositórios institucionais de acesso aberto à informação científica em bibliotecas de pesquisa**. In: Implantação e gestão de repositórios institucionais: políticas, memória, livre acesso e preservação. EDUFBA, 2010. p. 163-202.
- DE SOUZA, Arthur Heleno Lima Rodrigues et al. **O modelo de referência OAIS e a preservação digital distribuída**. Ciência da Informação, v. 41, n. 1, 2012.
- DILLO, Ingrid. LEEUW, Lisa de. **CoreTrustSeal**. Mitteilungen der Vereinigung Österreichischer Bibliothekarinnen & Bibliothekare, 2018, vol. 71, n. 1, pp. 162-170.
- MAIA, Cleusa Maria de Souza. FERRER, Igor Dias. BARROS, Diego Bil Silva. **Auditoria de repositórios digitais preserváveis**. Revista Ibero-Americana de Ciência da Informação, 2018.
- MASSON, Sílvia Mendes. **Os Repositórios digitais no âmbito da Sociedade Informacional**. Prisma. com, n. 7, p. 105-152, 2008.
- PAMPEL, Heinz. et al. **Making Research Data Repositories Visible: The re3data.org Registry**. Data Analysis and Knowledge Discovery, v. 30, n. 3, p. 26-34, 2014.
- SAYÃO, Luis Fernando; SALES, Luana Farias. **Algumas considerações sobre os repositórios digitais de dados de pesquisa**. Inf. Inf., Londrina, v. 21, n. 2, p. 90-115, 2016.

### AGRADECIMENTOS

A FUNDAÇÃO ARAUCÁRIA PELA BOLSA DE EXTENSÃO PIBIS.